View Dependent Surface Material Recognition

Stanislav Mikeš and Michal Haindl[(✉)] [iD]

The Institute of Information Theory and Automation
of the Czech Academy of Sciences, Prague, Czechia
{xaos,haindl}@utia.cz
http://www.utia.cz/

Abstract. The paper presents a detailed study of surface material recognition dependence on the illumination and viewing conditions which is a hard challenge in a realistic scene interpretation. The results document sharp classification accuracy decrease when using usual texture recognition approach, i.e., small learning set size and the vertical viewing and illumination angle which is a very inadequate representation of the enormous material appearance variability. The visual appearance of materials is considered in the state-of-the-art Bidirectional Texture Function (BTF) representation and measured using the upper-end BTF gonioreflectometer. The materials in this study are sixty-five different wood species. The supervised material recognition uses the shallow convolutional neural network (CNN) for the error analysis of angular dependency. We propose a Gaussian mixture model-based method for robust material segmentation.

1 Introduction

The visual appearance of surface materials and object shapes are crucial for visual scene understanding or interpretation. Visual aspects of surface materials which manifest themselves as visual textures even if there is still missing a rigorous definition of the texture [6]. Thus reliable visual scene interpretation cannot avoid a sound texture recognition quality. The correct recognition is hindered by the considerable variability of a material appearance and thus its corresponding textural representation based on changing observation conditions. Numerous texture recognition methods were published but we are not aware of a method which accounts for simultaneously variable illumination and viewing angle.

Most materials classification methods which respect illumination and view changes are restricted to BRDF (Bidirectional Reflectance Distribution Function) material representation [10,15,17] which neglects not only self-occlusion and inter-reflection material properties but also the essential spatial material features. The per-pixel SVM (Support Vector Machine) classification [15] is based on spectral BRDF and detects from training samples a discriminative illumination. The paper [10] is restricted to 5 steel classes, BRDF, and illumination changes only. [25] demonstrate the usefulness of several images of a material sample with different view-light conditions for material identification. [11] studies nine fabric classes with 2000 samples recognition using SIFT (Scale Invariant Feature Transform) and CNN (Convolutional Neural Network) features from albedo images and three concatenated normals into a single image. The authors

[23] used convolutional neural network for 12 material classes (with 100 images per class) recognition represented in 4D light-field measurements and achieved 7% accuracy improvement compared to single view images only. Material BRDF estimation from the 4D light-field using a convolutional neural network is also studied in [17]. The BTF classification exceptions are the papers [16] which studies the effects of illumination patterns, rotation, moreover, the scale for bag-of-words classification and [24] which uses an SVM classifier for synthetically generated material samples in the BTF representation. For detailed evaluation of ten previously published texture segmenters on BTF dataset see mosaic.utia.cas.cz [5].

Real surface material visual appearance is a very complex physical phenomenon which intricately depends on the incident and reflected spherical angles, time, and light spectrum among other physical quantities. The general and physically correct material reflectance function should be at least sixteen dimensional [6] which is recently unmeasurable, and even if some simplifying assumptions have to be inevitably accepted, the essential dependencies have to be respected. Among them, these are spectral, illumination, and viewing parameters.

Our main contributions are:

– Introduction of a new large BTF wood database measurement with 65 wood species.
– A first detailed evaluation of material recognition accuracy dependent on view and illumination angles.
– A detailed recognition accuracy evaluation of the BTF wooden mosaics mapped on generated 3D surfaces and using different viewing and illumination learning subsets.

1.1 Bidirectional Texture Function

The seven-dimensional bidirectional texture function (BTF) reflectance model is currently the state-of-the-art general reflectance function model approximation which can be simultaneously measured and modeled [3,6].

Multispectral BTF is a seven-dimensional function, which considers measurement dependency on color spectrum, planar material position, as well as its dependence on illumination $_{(i)}$ and viewing $_{(v)}$ angles: $BTF(r, \theta_i, \phi_i, \theta_v, \phi_v)$, where the multiindex $r = [r_1, r_2, r_3]$ specifies planar horizontal and vertical position in material sample image, r_3 is the spectral index and θ, ϕ are elevation and azimuthal angles of the illumination and view direction vectors. The BTF measurements comprise a whole hemisphere of light and camera positions in observed material sample coordinates according to selected quantization steps.

2 BTF Wood Measurement Database

Since the accurate and reliable BTF acquisition is not a trivial task, only a few BTF measurement systems exist [3,6,12,18,19,21,22]. Several material databases were already published [2,9]. A BTF database of 84 measured materials in seven categories, each containing 12 different material samples, was presented in [24]. [16] measured in 150 lighting conditions eight categories with 90 samples in total in their dome system.

Table 1. Measured illumination and viewing elevation (θ) and azimuthal ϕ angles for six viewing angular sets A–F, the number of directions in each set and their relative weights in the whole BTF space.

Set	$\theta_{i,v}$	$\triangle\phi_{i,v}$	♯directions	Weight [%]
A	0	–	1	1
B	15	60	6	7
C	30	30	12	15
D	45	20	18	22
E	60	18	20	25
F	75	15	24	30
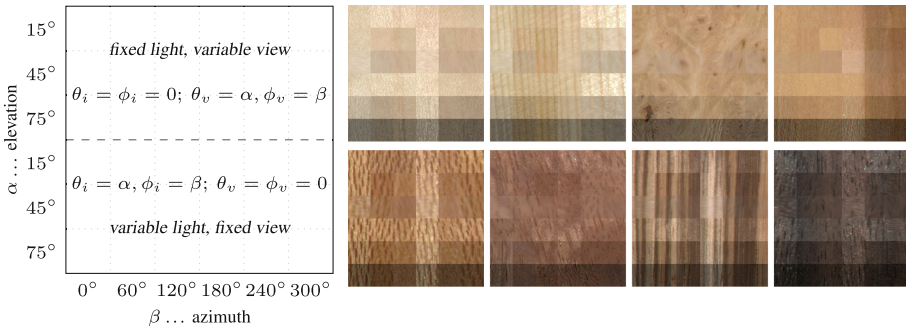
Fig. 1. Selected combinations (right) of illumination and viewing elevation (θ) and azimuthal ϕ measurement cutouts arranged as denoted in the left table from (rightwards, top-down) maple, spruce, oak, plum, sycamore, bubinga pomele, zebrano, and macassar, respectively.

We used a high precision robotic gonioreflectometer. The setup consists of independently controlled arms with camera and light. Its parameters such as angular precision 0.03°, spatial resolution 1000 DPI, or selective spatial measurement, classify this gonioreflectometer to the state-of-the-art devices. The typical resolution of the area of interest is around 2000×2000 pixels, sample size 7×7 [cm], sensor distance ≈ 2 [m] with field of view angle 8.25° and each of them is represented using at least 16-bit floating point value for a reasonable representation of high-dynamic-range visual information. Illumination source are eleven LED array, each having flux 280 lm at 0.7 A, spectral wavelength $450-700$ [nm], and have its optics. The memory requirements for storage of single material sample amount to 360 gigabytes per color channel but can be much more for a more precise spectral measurement.

We measured each wood sample in 81 viewing positions n_v and 81 illumination positions n_i resulting in 6561 images per sample (4 tera-bytes of data). Table 1 summarizes for each combination of elevation and azimuthal angle the number of measurements (♯) and the corresponding angular difference steps ($\triangle\phi_{i,v}$). The spatial resolution of the rectified original measurements was $M \times N \approx 1800 \times 1800$ pixels (Fig. 1).

The BTF wood database, used in this study, contains veneers from sixty-five varied European, African, and American wood species. The wood species are acacia tree, afzelia, alder, anigre, apple tree, ayous, cedar, elm, etimoe, eucalyptus, fir, gabon, hornbeam, iroko, jatoba, larch, limba, linden, macassar, mahogany, mansonia, meranti, merbau, movingui, olive tree, ovangkol, padouk, pear, pine, plane tree, plum, satin, teak, tineo, tulipwood, wenge, zebrawood; two species of cherry trees, birchs, ash trees, bamboo, bubingas, palisanders, spruces, and beeches; three species of oaks and walnuts; four species of maple trees.

3   Convolutional Neural Network

We use the open source TensorFlow library [1] for implementation of the convolutional neural network (CNN) for our experiments either as the classifier or as the source of features [13,14,20]. We used the shallow network with three convolution layers, 64 5×5 kernels, ReLU (rectified linear unit) activation functions, max pooling with stride 2, and local response normalization (for details see [14]). Another two fully connected layers (384 and 192 units) use the rectified linear activation, and finally, the linear transformation is applied to produce logits. CNN accept input image of the size with 48×48 in three spectral bands (RGB pixels). The images (64×64 patches) are processed as follows: they are cropped to 48×48 pixels, centrally for evaluation or randomly for training; they are per image standardized to make the model insensitive to dynamic range. For training, we additionally apply a series of random distortions to increase the dataset size artificially. These are image flip from left to right (H), image flip from up to down (V), rotation about $90°$ (R), and distortion of brightness and contrast. The model marking ($\mathbf{X}_{\mathrm{HVR}}^{\mathrm{d}}$) shows these training options (V, H, R). The network is trained to perform N-way classification using multinomial logistic regression. For regularization, we apply the common l_2 loss losses to all learned variables. The objective function for the model is the cross-entropy loss plus L_2 loss. For training this model we use the standard gradient descent optimizer with a learning rate (starting from 0.1) that exponential decay (0.1) over time (350 epochs), and random weights initialization.

4   Mixture-Based Segmenter

We propose the segmenter based on the Gaussian mixture model similar to the unsupervised algorithm in [4], but their original Markovian textural features we have replaced with the parameters (Θ) obtained from the learned CNN model on the floating window. Thus the segmenter is invariant to both illumination and viewing angles changes and benefits from strong noise suppression property of the Gaussian mixture model. The number of components (K) is variable based on the Kullback Leibler divergence estimation. The Gaussian mixture model for CNN parametric representation is

$$p(\Theta_{x,y}) = \sum_{i=1}^{K} p_i \, p(\Theta_{x,y} \mid \mu_i, \Sigma_i), \tag{1}$$

$$p(\Theta_{x,y} \mid \mu_i, \Sigma_i) = (2\pi)^{\frac{-n}{2}} |\Sigma_i|^{-\frac{1}{2}} \exp \left\{ -\frac{(\Theta_{x,y} - \mu_i)^T \Sigma_i^{-1} (\Theta_{x,y} - \mu_i)}{2} \right\}, \tag{2}$$

where x, y spatial coordinates, μ, Σ the Gaussian data model parameters, and $n = 66+3+2$ the number of features (CNN parameters, local color, and spatial coordinates). The mixture model Eqs. (1), (2) are solved using the EM algorithm and the parametric vectors representing texture mosaic pixels are assigned to the classes according to the highest component probabilities.

5   Results

All our experiments were provided on two BTF wooden sets. The first set contains $426465 = 65 \times 81^2$ measured wooden textures from 65 different wood species, while the second set contains ten synthetic BTF wooden mosaics from the Prague texture segmentation data-generator and benchmark [5,7].

5.1   Learning from the Complete BTF Space

Each measured BTF space image was divided into four quadrants. From each quadrant randomly selected 64×64 cutouts for every single combination of an illumination and viewing angle were chosen, i.e., **a** from left upper, **b** from lower left, **c** from upper right, and **d** from lower right. Figures 2 and 3 show how the CNN recognition accuracy depends on the viewing elevation angle for different learning scenarios averaged over all wooden species and all illumination angles. The short line segments in right in these graphs (Figs. 2 and 3) highlight the corresponding weighted average value between angles $0°-75°$.
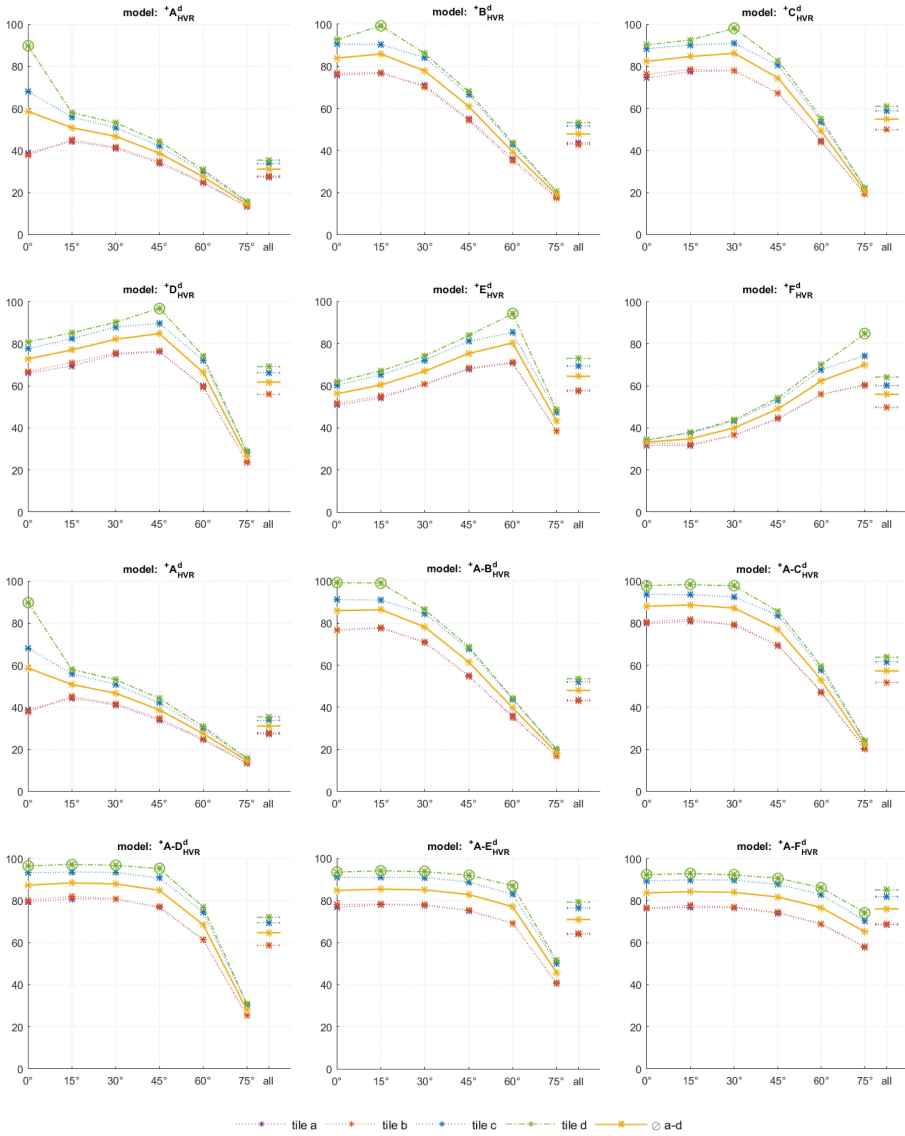
Viewing Angle Dependence and Horizontal Learning Flip. The CNN model $(\mathbf{X}_H^b)$ was learned on the cutout **b** in all classification experiments in cutouts **a–d**. The accuracy curves (Fig. 2) show the similar ranking in all eleven angle combination experiments the best hold out results are on the **a** cutous (always in the direction along the grain) then **d** (close parallel with **b**) and finally **c** (distant parallel with **b**). The best classification accuracy is reached if all sets are used for learning $(\mathbf{A}\text{-}\mathbf{F}_H^b)$. Table 2-left show the weighted (Table 1) averaged accuracy for all four possible cutouts and viewing elevation ranges from $0°$ until $0°-75°$. This table confirms the best possible results for the resubstitution accuracy estimates (**b**) for every angular range and the overall best accuracy (85.6) for the A–F range, followed with results on the **a** cutout. The worst accuracy has the most distant cutout **c** with the exception on the range A.

Table 2. Weighted averaged accuracy for the **a–d** cutouts over viewing elevation (θ) and azimuthal (ϕ) angles. CNN learning cutout is **b/d**.

	$\mathbf{X}_H^b$						$\mathbf{X}_{HVR}^d$					
θ_v [°]	0	0–15	0–30	0–45	0–60	0–75	0	0–15	0–30	0–45	0–60	0–75
a	36.3	51.7	59.9	67.9	75.1	79.8	27.4	43.3	51.8	58.6	64.0	68.5
b	38.2	54.5	63.6	72.6	80.2	85.6	27.9	43.0	51.9	58.8	64.2	68.8
c	29.2	44.0	50.9	57.2	63.5	67.5	33.8	52.1	61.7	69.5	76.4	81.8
d	29.1	44.7	52.4	58.8	65.1	69.4	35.4	53.5	63.8	72.0	79.2	85.0
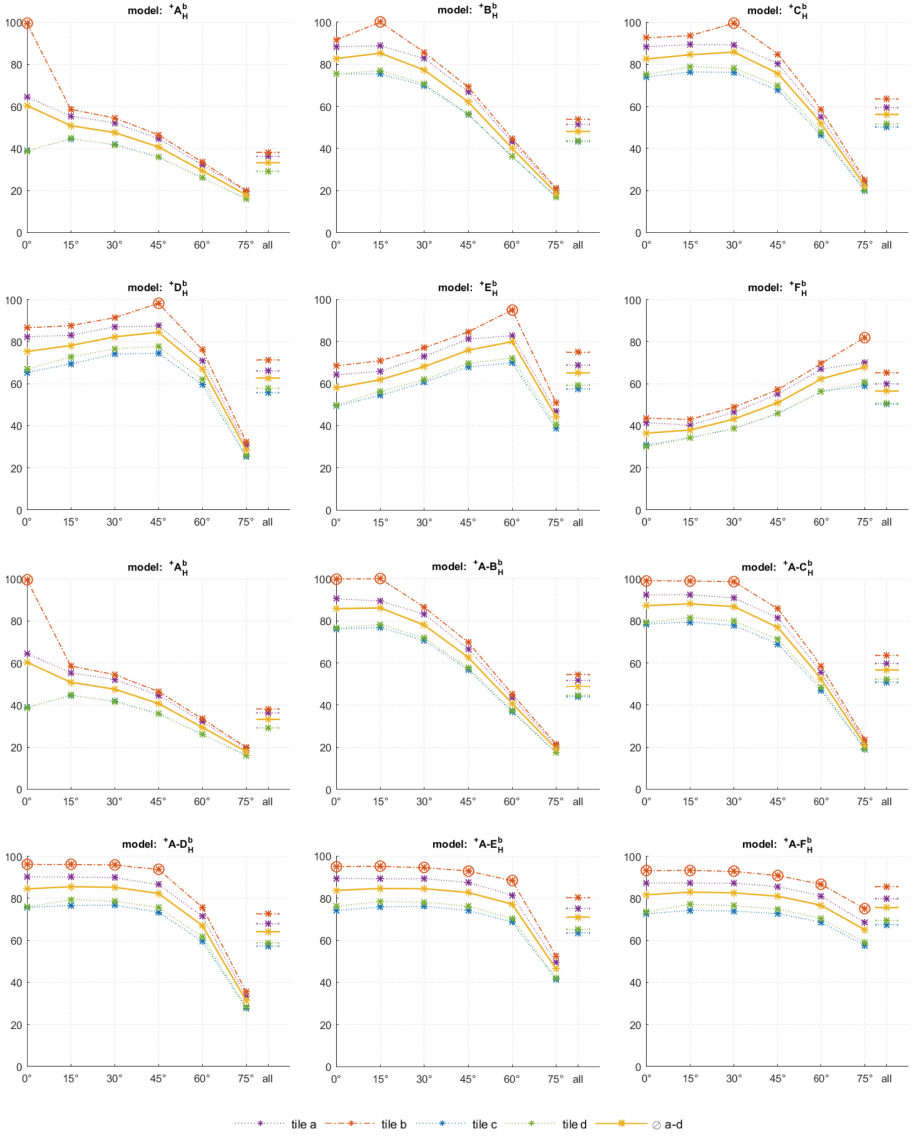
Fig. 2. Classification accuracy of the model $\mathbf{X}_H^b$ for cutouts a, b, c, d, a–d after learning from the b cutouts (⊛) with the horizontal flip obtained from the set $X \in \{A, B, C, D, E, F\}$ (see Table 1). The vertical axes show accuracy in %, while the horizontal axes show the viewing elevation angles.

Fig. 3. Classification accuracy of the model $\mathbf{X}_{HVR}^d$ for cutouts a, b, c, d, a–d after learning from the d cuts (⊛) with the horizontal, vertical flips and 90° rotation. The vertical axes show accuracy in %, while the horizontal axes show the viewing elevation angles.

Viewing Angle Dependence with Horizontal, Vertical Flip and Sample Rotation Learning. The CNN model ($\mathbf{X}_{\text{HVR}}^{\text{d}}$) was learned on the cutout **d** in all classification experiments in Fig. 3. The accuracy curves show the similar ranking again in all eleven experiments the best results are on the **c** cutous (in the direction along the grain), effects on **a, b** cutouts are very similar. The best classification accuracy is reached if all sets are used for learning ($\mathbf{A}$-$\mathbf{F}_{\text{HVR}}^{\text{d}}$). Table 2-right show the weighted averaged accuracy for all four possible cutouts and viewing elevation ranges from $0°$ until $0° - 75°$. As expected, also in this table the best possible results are the resubstitution estimates (**d**) for every angular range and the overall best accuracy is 85.0 for the A–F range, very similar with the corresponding result in the Table 2-left. The results on the cutouts **a, b** are very similar for all tested angular ranges. This is the consequence of the additional learning data (vertical flip and rotation). Comparing the results in both tables, it is possible to see that the additional flip and rotation learning data are advantageous only for the closest cutout (**c**) for $\mathbf{X}_{\text{HVR}}^{\text{d}}$ and larger angular ranges $> \langle 0; 30 \rangle$ but not for all ranges in the more distant cutouts (**a**) and all experiments on the A set only. However, the accuracy difference is very small for comparable cutouts and angles mostly in the range of 1%. Poor performance on the A set in both part of Table 2 illustrates inadequate learning set size (one image only) and the perpendicular viewing and illumination angle which is the standard but a very insufficient representation of the huge material appearance variability. The difference between the best and the worst accuracy in each angular range is gradually increasing from set A until A–F for both tables. This accuracy difference is between 38.3% (Table 2-left-**c**) and 49.6% (Table 2-right-**d**).

5.2 Synthetic BTF Wooden Mosaic

This experiment used the synthetic BTF wooden mosaic scenes from the Prague texture segmentation data-generator and benchmark [5,7] as well as the online evaluation capability (Table 3) of this web-based (http://mosaic.utia.cas.cz) service. The benchmark ranks segmentation experiments or algorithms results according to a chosen criterion. The benchmark has implemented twenty-seven frequented evaluation criteria (see [5] for their detailed explanation) categorized into region-based, pixel-wise, clustering comparison criteria, and consistency measures criteria sub-groups.

The benchmark test 1024×1024 mosaics layouts and each cell texture membership are randomly generated and filled with the measured BTF wooden textures from the large UTIA BTF database [8]. The BTF wood measurements are mapped on the randomly generated 3D spline surfaces. Each surface region is mapped with a physically correct wood material measurement which precisely corresponds to the local illumination and viewing conditions, and as such it represents the state-of-the-art realistic material visual representation [6] and also the best available current texture segmentation benchmark. The benchmark allows generating an unlimited number of experimental physically correct mosaics with exactly known segmentation ground truth. All benchmark mosaic experiments were carried out with the Gaussian mixture-based supervised classifier (Sect. 4) applied to CNN learned features. Table 3 illustrates the differences between classification accuracy between models learned on different viewing angle training subsets.

Table 3. BTF wood benchmark results; (Benchmark criteria: CS = correct segmentation; OS = over-segmentation; US = under-segmentation; ME = missed error; NE = noise error; O = omission error; C = commission error; CA = class accuracy; CO = recall - correct assignment; CC = precision - object accuracy; I. = type I error; II. = type II error; EA = mean class accuracy estimate; MS = mapping score; RM = root mean square proportion estimation error; CI = comparison index; GCE = Global Consistency Error; LCE = Local Consistency Error; dD = Van Dongen metric; dM = Mirkin metric; dVI = variation of information; $\bar{f}$ are the performance curves integrals); small numbers are the corresponding measure rank over the listed methods.

	Benchmark – BTF wood										
	A-D$_{HVR}^d$	B$_{HVR}^d$	E$_{HVR}^d$	A-C$_{HVR}^d$	F$_{HVR}^d$	C$_{HVR}^d$	D$_{HVR}^d$	A$_{HVR}^d$	A-F$_{HVR}^d$	A-B$_{HVR}^d$	A-E$_{HVR}^d$
	(2.90)	(3.05)	(3.90)	(3.95)	(5.90)	(6.00)	(6.71)	(7.48)	(8.33)	(8.48)	(9.24)
↑ CS	80.20^2	$\mathbf{81.02}^1$	80.13^3	77.87^5	78.72^4	71.83^8	72.97^7	77.01^6	70.27^{10}	69.70^{11}	71.30^9
↓ OS	50.63^3	56.30^5	58.36^8	47.98^2	51.44^4	58.80^9	$\mathbf{45.02}^1$	57.20^6	59.59^{10}	63.16^{11}	57.65^7
↓ US	1.28^2	4.72^5	3.31^3	7.19^9	4.18^4	5.13^7	5.23^8	7.74^{11}	$\mathbf{1.05}^1$	7.26^{10}	4.98^6
↓ ME	2.23^9	0.66^2	1.46^4	$\mathbf{0.58}^1$	2.74^{10}	1.16^3	1.90^7	4.44^{11}	2.23^8	1.89^6	1.76^5
↓ NE	3.80^9	$\mathbf{0.87}^1$	2.95^4	1.21^2	3.80^8	2.00^4	1.40^3	4.69^{11}	4.50^{10}	3.01^6	3.75^7
↓ O	4.39^3	3.06^2	5.88^7	5.10^6	10.39^{10}	4.41^4	$\mathbf{2.95}^1$	4.46^5	12.11^{11}	7.87^8	9.78^9
↓ C	$\mathbf{61.71}^1$	81.13^3	90.15^7	80.40^2	83.16^4	84.01^5	86.64^6	100.00^{10}	91.20^9	100.00^{11}	90.30^8
↑ CA	$\mathbf{87.12}^1$	86.75^3	86.82^2	85.96^4	84.47^5	83.24^7	82.60^8	83.79^6	81.47^{10}	81.48^9	80.54^{11}
↑ CO	$\mathbf{88.88}^1$	88.61^3	88.69^2	88.43^4	87.25^5	85.23^8	85.31^7	86.98^6	83.04^{10}	83.97^9	82.72^{11}
↑ CC	97.76^2	97.19^4	97.19^5	96.59^7	96.07^9	97.21^3	95.71^{10}	94.79^{11}	$\mathbf{97.77}^1$	96.70^6	96.56^8
↓ I.	$\mathbf{11.12}^1$	11.39^3	11.31^2	11.57^4	12.75^5	14.77^8	14.69^7	13.02^6	16.96^{10}	16.03^9	17.28^{11}
↓ II.	$\mathbf{0.25}^1$	0.32^4	0.38^5	0.44^7	0.48^{10}	0.32^3	0.47^8	0.67^{11}	0.28^2	0.47^9	0.38^6
↑ EA	91.74^2	91.35^3	$\mathbf{91.78}^1$	90.74^4	89.62^5	88.83^7	88.29^8	89.17^6	88.05^9	87.44^{10}	86.84^{11}
↑ MS	$\mathbf{87.63}^1$	87.26^3	87.42^2	86.65^4	85.17^5	83.83^7	83.46^8	84.56^6	81.89^{10}	82.09^9	81.01^{11}
↓ RM	1.77^4	1.94^6	1.52^2	1.71^3	$\mathbf{1.41}^1$	2.05^8	2.73^{11}	1.97^7	2.10^9	1.80^5	2.29^{10}
↑ CI	$\mathbf{92.47}^1$	92.07^3	92.32^2	91.39^4	90.43^5	89.85^7	89.33^8	89.93^6	89.17^9	88.66^{10}	88.12^{11}
↓ GCE	4.31^2	$\mathbf{4.27}^1$	4.67^3	5.00^7	6.62^{10}	4.75^4	4.78^5	6.90^{11}	4.98^6	5.17^8	6.03^9
↓ LCE	3.11^5	$\mathbf{2.38}^1$	3.42^8	2.78^3	3.39^7	2.74^2	2.82^4	3.59^9	3.81^{10}	3.18^6	4.00^{11}
↓ dD	6.70^2	$\mathbf{6.54}^1$	6.90^4	6.70^3	7.57^5	8.40^7	8.53^8	7.97^6	9.92^{10}	9.12^9	10.27^{11}
↓ dM	4.29^3	4.54^6	4.24^2	$\mathbf{3.94}^1$	4.47^5	5.18^7	6.20^9	4.41^4	6.80^{10}	5.90^8	6.97^{11}
↓ dVI	15.97^6	15.82^4	15.91^5	$\mathbf{15.61}^1$	15.81^3	16.39^8	16.03^7	15.70^2	16.73^{10}	16.48^9	16.74^{11}
↑ $\overline{CS}$	$\mathbf{74.65}^1$	74.11^2	72.42^3	72.37^4	70.53^6	68.17^7	68.05^8	69.03^5	63.00^{11}	65.24^9	63.07^{10}
↓ $\overline{OS}$	48.46^3	53.41^6	55.03^8	43.91^2	49.88^4	56.36^{10}	$\mathbf{43.00}^1$	53.48^7	56.35^9	58.68^{11}	53.10^5
↓ $\overline{US}$	$\mathbf{1.09}^1$	3.74^4	1.87^3	6.12^{10}	4.89^7	4.64^6	5.33^8	6.38^{11}	1.34^2	6.03^9	3.87^5
↓ $\overline{ME}$	8.01^5	$\mathbf{5.91}^1$	9.17^6	6.87^2	9.76^8	7.67^3	7.87^4	10.60^9	10.73^{10}	8.94^6	11.63^{11}
↓ $\overline{NE}$	8.73^5	$\mathbf{6.06}^1$	10.13^7	7.39^2	10.50^8	8.19^4	7.90^3	11.17^9	12.37^{10}	9.72^6	13.29^{11}
↑ $\overline{F}$	$\mathbf{92.27}^1$	91.87^3	92.17^2	91.21^4	90.20^5	89.56^7	89.04^8	89.71^6	88.85^9	88.31^{10}	87.75^{11}

The first row shows various view angle ranges (A, B, C, D, E, F, A–B, A–C, A–D, A–E, A–F) used for CNN learning as explained in Table 1. Small numbers are the corresponding criteria rank value in this table and the second row in Table 3 lists the average rank for the corresponding learning set. The best learning set is A–D (2×10^5 learning images with elevation angles $0°-45°$) with ten best performing criteria. This set is the best compromise between the learning variability and recognition accuracy.
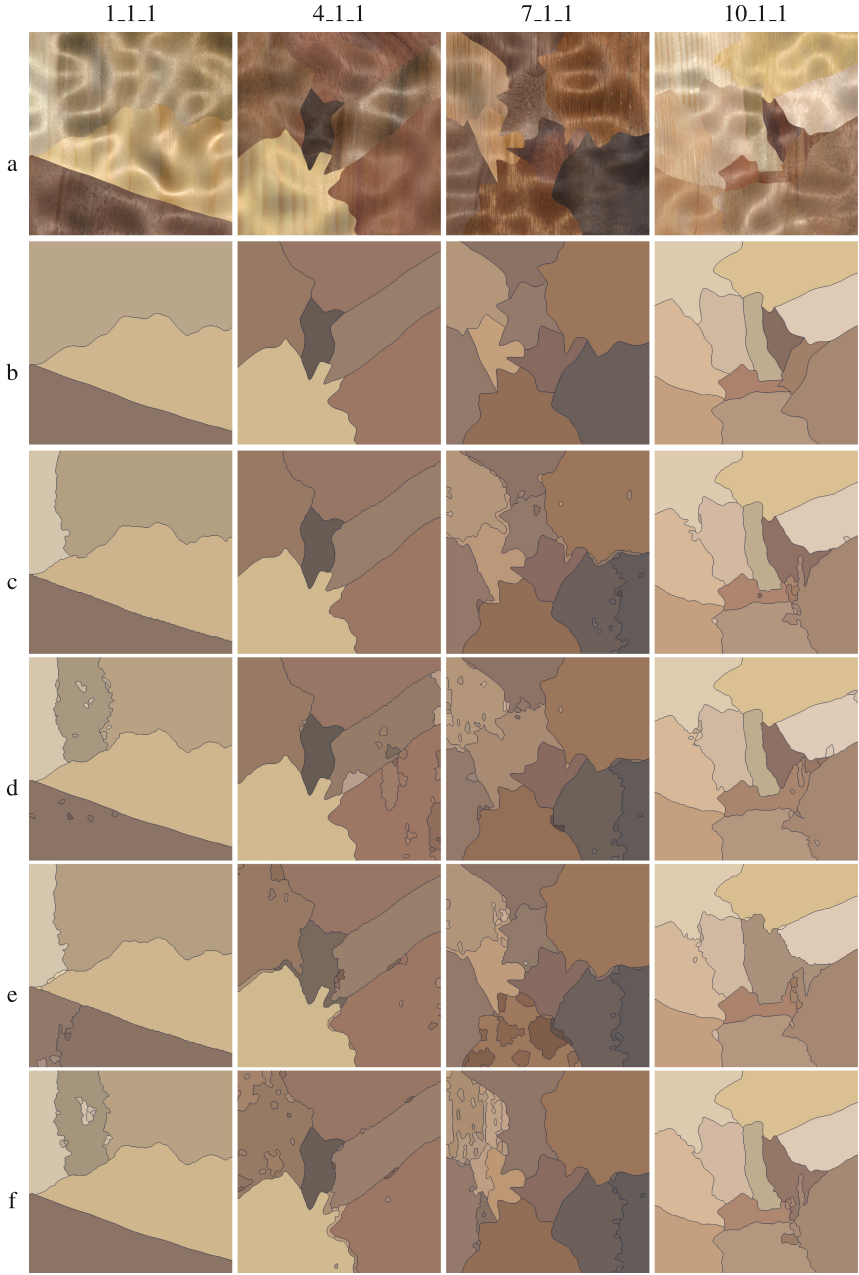
Fig. 4. Selected BTF benchmark mosaics (a), ground-truth (b), **A-D**$_{\text{HVR}}^{\text{d}}$ (c), **B**$_{\text{HVR}}^{\text{d}}$ (d), **E**$_{\text{HVR}}^{\text{d}}$ (e), **A-C**$_{\text{HVR}}^{\text{d}}$ (f) segmentation results, respectively.

The learning sets ordering in Table 3 is based on the best average rank over all segmentation criteria rightwards. The worst segmenter performance (eleven worst criteria) is learned from the set AE. This learning variant has the worst class accuracy, recall, and nine other criteria values, and none winning criterion.

The best-fixed learning elevation angle is $15°$ - the B set from the region-based criteria point of view, while the A set ($0°$) has the worst average rank and only the variation of information being the second best (Fig. 4).

6   Conclusions

This study presented two experiments of view and illumination-dependent material recognition analysis. The sixty-five wood species measured in the state-of-the-art bidirectional texture function representation is classified using the convolutional neural network. The novel Gaussian mixture based segmenter with CNN learned features is favourably evaluated on eleven different angular combinations. The results document sharp classification accuracy decrease when using standard texture recognition approach, i.e., small learning set size and the vertical viewing and illumination angle which is a very inadequate representation of the enormous material appearance variability. The ideal learning is to use the whole possible viewing angle range. The benchmark experiments suggest $15°$ to be the best single elevation angle and $0°-45°$ the best range of elevation angles. The mosaic surfaces are smooth without sharp declinations thus they possibly prefer narrower elevation ranges than an object in real visual scenes.

References

1. TensorFlow. Technical report, Google AI. http://www.tensorflow.org/
2. Dana, K.J., van Ginneken, B., Nayar, S.K., Koenderink, J.J.: Reflectance and texture of real world surfaces. Technical report CUCS-048-96, Columbia University, December 1996
3. Dana, K.J., Nayar, S.K., van Ginneken, B., Koenderink, J.J.: Reflectance and texture of real-world surfaces. In: CVPR, pp. 151–157. IEEE Computer Society (1997)
4. Haindl, M., Mikeš, S.: Unsupervised texture segmentation using multispectral modelling approach. In: Tang, Y., Wang, S., Yeung, D., Yan, H., Lorette, G. (eds.) Proceedings of the 18th International Conference on Pattern Recognition, ICPR 2006, vol. II, pp. 203–206. IEEE Computer Society, Los Alamitos, August 2006
5. Haindl, M., Mikeš, S.: Texture segmentation benchmark. In: Lovell, B., Laurendeau, D., Duin, R. (eds.) Proceedings of the 19th International Conference on Pattern Recognition, ICPR 2008, pp. 1–4. IEEE Computer Society, Los Alamitos, December 2008
6. Haindl, M., Filip, J.: Visual Texture. ACVPR. Springer, London (2013). https://doi.org/10.1007/978-1-4471-4902-6
7. Haindl, M., Mikeš, S.: A competition in unsupervised color image segmentation. Pattern Recogn. **57**(9), 136–151 (2016)
8. Haindl, M., Mikeš, S., Kudo, M.: Unsupervised surface reflectance field multi-segmenter. In: Azzopardi, G., Petkov, N. (eds.) CAIP 2015. LNCS, vol. 9256, pp. 261–273. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-23192-1_22

9. Hayman, E., Caputo, B., Fritz, M., Eklundh, J.-O.: On the significance of real-world conditions for material classification. In: Pajdla, T., Matas, J. (eds.) ECCV 2004. LNCS, vol. 3024, pp. 253–266. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-24673-2_21

10. Jehle, M., Sommer, C., Jähne, B.: Learning of optimal illumination for material classification. In: Goesele, M., Roth, S., Kuijper, A., Schiele, B., Schindler, K. (eds.) DAGM 2010. LNCS, vol. 6376, pp. 563–572. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15986-2_57

11. Kampouris, C., Zafeiriou, S., Ghosh, A., Malassiotis, S.: Fine-grained material classification using micro-geometry and reflectance. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9909, pp. 778–792. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46454-1_47

12. Koudelka, M.L., Magda, S., Belhumeur, P.N., Kriegman, D.J.: Acquisition, compression, and synthesis of bidirectional texture functions. In: Texture 2003: Third International Workshop on Texture Analysis and Synthesis, Nice, France, pp. 59–64, October 2003

13. Krizhevsky, A.: Learning multiple layers of features from tiny images. Master's thesis, University of Toronto, Canada (2009)

14. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)

15. Liu, C., Gu, J.: Discriminative illumination: per-pixel classification of raw materials based on optimal projections of spectral BRDF. IEEE Trans. Pattern Anal. Mach. Intell. **36**(1), 86–98 (2014)

16. Liu, C., Yang, G., Gu, J.: Learning discriminative illumination and filters for raw material classification with optimal projections of bidirectional texture functions. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2013

17. Lu, F., He, L., You, S., Chen, X., Hao, Z.: Identifying surface BRDF from a single 4-D light field image via deep neural network. IEEE J. Sel. Top. Sig. Process. **11**(7), 1047–1057 (2017)

18. Müller, G., Meseth, J., Sattler, M., Sarlette, R., Klein, R.: Acquisition, synthesis and rendering of bidirectional texture functions. In: Eurographics 2004, STAR - State of The Art Report, pp. 69–94. Eurographics Association (2004)

19. Ngan, A., Durand, F.: Statistical acquisition of texture appearance. In: Eurographics Symposium on Rendering. Eurographics (2006)

20. Pattanayak, S.: Pro Deep Learning with TensorFlow. Apress, New York (2017)

21. Sattler, M., Sarlette, R., Klein, R.: Efficient and realistic visualization of cloth. In: Eurographics Symposium on Rendering, June 2003

22. Wang, J., Dana, K.: Relief texture from specularities. IEEE Trans. Pattern Anal. Mach. Intell. **28**(3), 446–457 (2006)

23. Wang, T.-C., Zhu, J.-Y., Hiroaki, E., Chandraker, M., Efros, A.A., Ramamoorthi, R.: A 4D light-field dataset and CNN architectures for material recognition. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9907, pp. 121–138. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46487-9_8

24. Weinmann, M., Gall, J., Klein, R.: Material classification based on training data synthesized using a BTF database. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8691, pp. 156–171. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10578-9_11

25. Weinmann, M., Klein, R.: Material recognition for efficient acquisition of geometry and reflectance. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) ECCV 2014. LNCS, vol. 8927, pp. 321–333. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16199-0_23